

Self-adjoint Matrices are Equivariant

Michael Dellnitz*

*Department of Mathematics, Paderborn University, D-33095 Paderborn, Germany

January 26, 2017

Abstract

In this short note we prove that a matrix $A \in \mathbb{R}^{n,n}$ is self-adjoint if and only if it is equivariant with respect to the action of a group $\Gamma \subset \mathbf{O}(n)$ which is isomorphic to $\otimes_{k=1}^n \mathbf{Z}_2$. Moreover we discuss potential applications of this result, and we use it in particular for the approximation of higher order derivatives for smooth real valued functions of several variables.

Key words: self-adjoint matrix, equivariance, symmetry, Taylor expansion

AMS subject classifications. 15B57, 15A24, 37G40, 41A58

1 Introduction

Within this short note we prove a characterization for a matrix being *symmetric* – in the sense of $A = A^T$ – by using the notion of *equivariance*. The proof of this fact is not difficult at all, but to the best of the knowledge of the author so far the related result cannot explicitly be found in the literature.

However, in several articles concerning the development of dynamical systems for the solution of certain optimization problems this underlying equivariance structure is implicitly present (e.g. [1, 2, 3]), and one would expect that this is also the case in other applications. The point of this note is to state this characterization of $A = A^T$ explicitly, and this is done in Section 2. In Section 3 we discuss potential applications in equivariant bifurcation theory, and we illustrate concretely how this result can be used for the construction of simple approximations of derivatives of higher order for real valued functions.

2 Main Result

Let $\Sigma \subset \mathbf{O}(n)$ be the abelian group consisting of the 2^n matrices

$$\begin{pmatrix} \pm 1 & 0 & 0 & \cdots & 0 \\ 0 & \pm 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \pm 1 \end{pmatrix}.$$

Obviously for any diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}, \quad \lambda_j \in \mathbb{R}, \quad j = 1, 2, \dots, n,$$

we have

$$\sigma D = D \sigma \quad \forall \sigma \in \Sigma.$$

In fact, it is easy to verify that for an arbitrary matrix $B \in \mathbb{R}^{n,n}$ one has

$$\sigma B = B \sigma \quad \forall \sigma \in \Sigma \quad \Longleftrightarrow \quad B \text{ is a diagonal matrix.} \quad (1)$$

In this note we prove the following characterization:

Proposition 2.1. *A matrix $A \in \mathbb{R}^{n,n}$ is self-adjoint (i.e. $A = A^T$) if and only if there is an orthogonal matrix $V \in \mathbf{O}(n)$ such that*

$$\gamma A = A \gamma \quad \forall \gamma \in \Gamma, \quad (2)$$

where the group $\Gamma \subset \mathbf{O}(n)$ is defined by

$$\Gamma = \{V^T \sigma V : \sigma \in \Sigma\}.$$

Proof. Suppose that $A = A^T$. Then there is $V \in \mathbf{O}(n)$ such that

$$D = V A V^T$$

is a diagonal matrix. By (1) we have for all $\sigma \in \Sigma$

$$\sigma V A V^T = V A V^T \sigma \quad \Longleftrightarrow \quad V^T \sigma V A = A V^T \sigma V.$$

Therefore A satisfies the equivariance condition (2).

Now suppose that (2) is satisfied for some $V \in \mathbf{O}(n)$. Then the matrix $V A V^T$ commutes with every $\sigma \in \Sigma$, and by (1) it follows that $D = V A V^T$ is a diagonal matrix. Therefore

$$A^T = (V^T D V)^T = A$$

as desired. □

Remarks 2.2. (a) Observe that the implication " \implies " could also be proved by using the well know fact that two matrices A and B commute if there is an orthogonal transformation V such that both $V^T A V$ and $V^T B V$ are diagonal.

(b) By construction all the eigenvalues of every $\gamma \in \Gamma$ are 1 or -1 . In particular $\gamma^2 = I$ for all $\gamma \in \Gamma$. Moreover, by (a) the matrix A and all $\gamma \in \Gamma$ possess the same set of eigenvectors.

(c) Obviously analogous results can be obtained for Hermitian or normal matrices: Using essentially the same proof as in Proposition 2.1 one can show that a matrix $A \in \mathbb{C}^{n,n}$ is normal (i.e. $AA^* = A^*A$) if and only if there is a unitary matrix $W \in \mathbf{U}(n)$ such that

$$\gamma A = A\gamma \quad \forall \gamma \in \Gamma,$$

where the group $\Gamma \subset \mathbf{U}(n)$ is defined by

$$\Gamma = \{W^* \sigma W : \sigma \in \Sigma\}.$$

3 On Applications

Proposition 2.1 could be used to look at results for symmetric matrices in the light of the equivariance condition (2). For instance a result from [4] on the genericity of the structure of eigenspaces would imply the well known fact that generically eigenspaces of self-adjoint matrices are one-dimensional. (Simply observe that $\Gamma \cong \otimes_{k=1}^n \mathbf{Z}_2$ possesses only one-dimensional (absolutely) irreducible representations.)

A potentially more interesting application may be the analysis of symmetry breaking bifurcations for gradient systems since in this case the Jacobian would be equivariant according to (2). This could particularly be useful for bifurcation problems where the (symmetric) steady state solution does not depend on the bifurcation parameter. In fact, some time ago the author himself has co-authored an article on "equivariant (and) self-adjoint matrices" [5], and it could be interesting to reconsider these results by taking the insight provided by Proposition 2.1 into account.

However, within this note let us focus concretely on one implication involving Taylor expansions. In this context the following immediate consequence of Proposition 2.1 strongly indicates that the result could, for instance, be used to develop a novel general approach for the construction of higher order stencils for real valued functions of several variables.

Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is smooth in a neighborhood of $\bar{x} \in \mathbb{R}^n$. In the following we use Proposition 2.1 to construct a four-point-stencil which provides a second order approximation of evaluations of the fourth order derivative in \bar{x} . For convenience we write the Taylor expansion of f in \bar{x} as

$$f(\bar{x} + h) = f(\bar{x}) + \nabla f(\bar{x})^T h + \frac{1}{2} h^T H(\bar{x}) h + \sum_{j=3}^{\infty} g_j(\bar{x}, h),$$

where $g_j(\bar{x}, h) = O(\|h\|^j)$, $j = 3, 4, \dots$, and $H(\bar{x})$ is the Hessian matrix of f at \bar{x} .

Corollary 3.1. *Denote by $\Gamma(\bar{x})$ the group in Proposition 2.1 corresponding to the Hessian matrix $H(\bar{x})$. Then for all $\gamma \in \Gamma(\bar{x})$ we have*

$$f(\bar{x} + \gamma h) - 2f(\bar{x}) + f(\bar{x} - \gamma h) = h^T H(\bar{x}) h + 2g_4(\bar{x}, \gamma h) + O(\|h\|^6), \quad (3)$$

and therefore for all $\gamma_1, \gamma_2 \in \Gamma(\bar{x})$

$$\begin{aligned} & f(\bar{x} + \gamma_1 h) + f(\bar{x} - \gamma_1 h) - f(\bar{x} + \gamma_2 h) - f(\bar{x} - \gamma_2 h) = \\ & = 2(g_4(\bar{x}, \gamma_1 h) - g_4(\bar{x}, \gamma_2 h)) + O(\|h\|^6). \end{aligned} \quad (4)$$

In particular, $f(\bar{x} + \gamma_1 h) + f(\bar{x} - \gamma_1 h) - f(\bar{x} + \gamma_2 h) - f(\bar{x} - \gamma_2 h) = O(\|h\|^4)$.

Proof. For $h \in \mathbb{R}^n$ and $\gamma_j \in \Gamma(\bar{x})$ ($j = 1, 2$) we compute using (2) and the fact that $\Gamma(\bar{x}) \subset \mathbf{O}(n)$

$$f(\bar{x} \pm \gamma_j h) = f(\bar{x}) \pm \nabla f(\bar{x})^T \gamma_j h + \frac{1}{2} h^T H(\bar{x}) h \pm g_3(\bar{x}, \gamma_j h) + g_4(\bar{x}, \gamma_j h) \pm g_5(\bar{x}, \gamma_j h) + \dots$$

Therefore

$$\begin{aligned} f(\bar{x} + \gamma_1 h) + f(\bar{x} - \gamma_1 h) &= 2 \left(f(\bar{x}) + \frac{1}{2} h^T H(\bar{x}) h + g_4(\bar{x}, \gamma_1 h) + O(\|h\|^6) \right) \\ f(\bar{x} + \gamma_2 h) + f(\bar{x} - \gamma_2 h) &= 2 \left(f(\bar{x}) + \frac{1}{2} h^T H(\bar{x}) h + g_4(\bar{x}, \gamma_2 h) + O(\|h\|^6) \right), \end{aligned}$$

and (3), (4) immediately follow. \square

Obviously, if $\gamma_1 = \pm \gamma_2$ then this result is not useful. However, for all other choices of γ_j this leads to interesting approximations of the fourth order derivative as long as h is not an eigenvector of γ_j ($j = 1, 2$).

Example 3.2. Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ be defined by

$$f(x_1, x_2, x_3) = x_1 x_2 x_3^2 + x_1^2 - 3x_2^2 + x_2 \sin(x_1) - x_2^2 x_3^2.$$

We choose $\bar{x} = (1, 1, 1)^T$ and compute

$$H(\bar{x}) = \begin{pmatrix} 2 - \sin(1) & 1 + \cos(1) & 2 \\ 1 + \cos(1) & -8 & -2 \\ 2 & -2 & 0 \end{pmatrix}.$$

The choice of

$$\sigma_1 = I \quad \text{and} \quad \sigma_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

leads to

$$\gamma_1 = I \quad \text{and} \quad \gamma_2 = \begin{pmatrix} 0.9225 & 0.3723 & 0.1015 \\ 0.3723 & -0.7896 & -0.4877 \\ 0.1015 & -0.4877 & 0.8671 \end{pmatrix}.$$

For $h = (0.2, 0.05, 0.1)^T$ we obtain

$$f(\bar{x} + h) + f(\bar{x} - h) - f(\bar{x} + \gamma_2 h) - f(\bar{x} - \gamma_2 h) \approx 6.40 \cdot 10^{-5},$$

and for $h = \frac{1}{10}(0.2, 0.05, 0.1)^T$ one computes

$$f(\bar{x} + h) + f(\bar{x} - h) - f(\bar{x} + \gamma_2 h) - f(\bar{x} - \gamma_2 h) \approx 6.38 \cdot 10^{-9}$$

as expected.

References

- [1] P. Schönemann. On two-sided orthogonal Procrustes problems. *Psychometrika*, 33(1):19–33, 1968.
- [2] R.W. Brockett. Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems. In *Decision and Control, 1988., Proceedings of the 27th IEEE Conference on*, pages 799–803. IEEE, 1988.
- [3] R.W. Brockett. Least squares matching problems. *Linear Algebra and its Applications*, 122:761–777, 1989.
- [4] M. Golubitsky, I. Stewart, and D. Schaeffer. *Singularities and Groups in Bifurcation Theory*. Springer, 1988.
- [5] M. Dellnitz and I. Melbourne. Generic movement of eigenvalues for equivariant self-adjoint matrices. *Journal of Computational and Applied Mathematics*, 55(3):249–259, 1994.